

Article

# Efficient Calibration of a Conceptual Hydrological Model Based on the Enhanced Gauss–Levenberg–Marquardt Procedure

Andrej Vidmar \*, Mitja Brilly, Klaudija Sapač  and Andrej Kryžanowski

Faculty of Civil and Geodetic Engineering, University of Ljubljana, Jamova 2, 1000 Ljubljana, Slovenia; mitja.brilly@fgg.uni-lj.si (M.B.); klaudija.sapac@fgg.uni-lj.si (K.S.); andrej.kryzanowski@fgg.uni-lj.si (A.K.)

\* Correspondence: andrej.vidmar@fgg.uni-lj.si; Tel.: +386-1-42-54-052

Received: 4 May 2020; Accepted: 29 May 2020; Published: 31 May 2020



**Abstract:** Various models were developed in the past to simulate different hydrological processes. However, discrepancies between simulated and observed values are still significant and pose a challenge to many researchers. Models contain many parameters that cannot be directly measured. The values of most of these parameters are determined in the calibration process conditioning the efficiency of such models. This paper introduces the use of the enhanced Gauss–Levenberg–Marquardt (GLM) procedure in combination with the singular value decomposition (SVD) and Tikhonov regularization to improve the process of hydrological model calibration. The procedure is tested on a freely available hydrological model using a synthetic dataset. Based on several efficiency measures, the GLM procedure, in combination with SVD and Tikhonov regularization, was found to provide efficient model history matching and almost perfect parameter calibration. Moreover, by comparing the results of the proposed procedure with the results of global evolutionary calibration procedures, it was found that the only calibration using the combined GLM procedure gave a perfect fit in low flows. Last but not least, the noise in the calculation results with the combined GLM method was practically the same in either the calibration or validation procedure, suggesting that only computational noise remained in the results.

**Keywords:** calibration model efficiency; Parameter Estimation Tool (PEST); HBV-light; Gauss–Levenberg–Marquardt procedure; SVD; Tikhonov regularization; parallel computing

## 1. Introduction

Simulation of hydrological processes using computer models has a long and rich history [1,2]. Hydrological models are merely a mathematical approximation of a variety of dynamic processes in highly heterogeneous conditions in nature. A model's variables are spatially heterogeneously distributed, while we have limited datasets available. The same is true for data on how water moves over and below the surface [3–6]. Different hydrological processes dominate different river catchments due to natural conditions. Therefore, various models were developed to simulate these hydrological processes more or less successfully [1,2,7,8]. More than 60 different hydrological models are available for rainfall–runoff simulations [1]. Due to noise, the discrepancies between simulation results and measurements are still significant and challenging for many researchers [6,9–16]. One of the reasons for the discrepancy between measurements and results of simulations can also be attributed to a lack of knowledge [17].

Discrepancies between observed data and simulated values, i.e., noise, can be produced by many sources. First, the data we use in the model alone is a source of the so-called primary noise resulting from measurements and observations [6,10,18–25]. Second, the source of noise is the structure of the

model itself, i.e., structural noise. The structural noise of modeling is the fact that the model is not a duplicate of the real state of nature. The structural noise of the modeling can be further divided into conceptual noise [26], model design noise, and model calibration noise. Conceptual noise is caused by a set of selected equations of complex hydrological processes that simulate natural phenomena [25,27]. For example, when highly nonlinear phenomena are modeled using linear equations, or due to the lack of data, certain parts of the hydrological cycle can be easily extracted from the model producing structural noise [25,26]. The design noise of the model is conditioned by the spatial and temporal discretization of the model when preparing the input data, i.e., with the same concept, the same river catchment can be designed differently. A model's calibration noise is conditioned by the choice of the calibration method by which the model parameters are calibrated.

In practice, hydrological modeling involves various models, stochastic or deterministic and conceptual or empirical. Hydrological models contain many parameters that are integral to the model's equations and cannot be determined directly by field measurements. At best, only the limits, within which the parameters take value according to their physical meaning or practical experience, can be defined. The parameters are determined in the calibration procedure when the calculated results are compared with the measurements [28–30]. The question of model calibration is as old as modeling itself, and several different calibration methods have evolved.

The oldest procedure for model calibration is the trial-and-error method, where the modeler manually changes the parameter values and visually analyzes the results. This approach was also used in the original design of the HBV (Hydrologiska Byråns Vattenavdelning) hydrological model [31,32]. The process is entirely subjective and conditioned by the knowledge and experience of the modeler and is mainly time-consuming. However, the modeler gets a direct sense of the sensitivity of the model and the importance of the individual parameters.

Over time, many computer-based methods for model calibration have evolved which can otherwise be used for a variety of purposes and in different scientific areas: the Monte Carlo method [33,34], the general probability of estimation uncertainty [35], the artificial neural network [36], the genetic algorithm (GAP) [18], the hybrid harmony search algorithm [37], the covariance matrix adaptation evolution strategy (CMA-ES) global optimization scheme [38–40], and the generalized reduced gradient [41]. Different theoretical bases give different calibration processes, where we generally distinguish between Newton-type methods, based on the first and second derivatives of the objective function, and global evolved methods, based on a stochastic approach [42]. As a disadvantage of Newton-type procedures, the literature reports about the phenomenon that the process is lost at the local minimum and is unable to find a possibly better global solution [42].

The calibration of complex nonlinear conceptual models has been the subject of intense studies for decades [17]. Recently, the popularity of the process of finding the inverse solution of the equations of a deterministic conceptual model, using a combination of Gauss–Levenberg–Marquardt (GLM) transformation, singular value decomposition (SVD), and Tikhonov regularization, is increasing [26,43–45]. The GLM method has been successfully used in the calibration of groundwater models [46–48], models of biophysical processes in agricultural systems [49], the optimal power flow (OPF) models to determine the best operating levels for electric power plants [50], petroleum and geothermal reservoir models [51], the estimation of soil loss model [52], and other complex environmental models [53]. However, to the authors' knowledge, the successful application of the GLM method in combination with SVD and Tikhonov regularization in the calibration of a conceptual hydrological model has not been reported yet by other researchers. Therefore, this paper aims to demonstrate that the combination of processes mentioned can provide similar or even better calibration results compared to other established approaches for hydrological model calibration.

The main objectives of this paper are to (1) present and apply the multicriteria GLM procedure in combination with singular value decomposition and Tikhonov regularization to calibrate the hydrological model, (2) use the existing freely available hydrological model with synthetic runoff data [18,54,55], which is required to determine the appropriate parameter settings of calibration to

reduce calibration noise and to improve the calibration efficiency, (3) compare the performance of the GLM procedure in combination with SVD and Tikhonov regularization with global evolutionary calibration procedures, and (4) accelerate the calibration process using parallel computing. Moreover, the reader can find a detailed description of the proposed procedure in Supplementary Materials.

## 2. Materials and Methods

In the section below, we present the Gauss–Levenberg–Marquardt procedure, singular value decomposition, and Tikhonov regularization, which is used in the inverse solutions of the equations of a deterministic conceptual model. In addition, the hydrological model and data are described. A methodology for model calibration and validation is presented. All procedures, as mentioned earlier, i.e., GLM, SVD, and Tikhonov regularization, are included in the PEST software [53]. Therefore, for this study, we used the PEST software tool, which is an industry-standard software package designed to evaluate parameters and analyze the uncertainties of complex environmental and other computer models [53]. Specifically, we used the PEST\_HP calibration tool (HP means very parallelized), which is equal to PEST [56,57] but PEST\_HP uses parallel computation, resulting in faster calculation performance.

### 2.1. Gauss–Levenberg–Marquardt Method

The Gauss–Levenberg–Marquardt method (GLM) is a blend of gradient descent and Gauss–Newton methods. In the literature, it is also referred to as the “Levenberg–Marquardt” algorithm (LMA or just LM), or the “damped least-squares” method (DLS) [58]. The parameter estimation is based on Equation (1). When the target function is far from being the smallest, the GLM operates by the steepest gradient method, but around the minimum, the GLM begins to act as a Gauss–Newton method. The Marquardt lambda in Equation (1) is a blending factor, which determines the mix between the gradient descent and the Gauss–Newton methods.  $\lambda$  is the so-called “Marquardt parameter,” or synthetic “Marquardt lambda,” named after Marquardt [58], who introduced this methodology. However, the use of this parameter was pioneered by Levenberg [59]. If the error is increasing, the function is far from its minimum and lambda should be increased approaching the gradient descent method. On the other hand, if the error is decreasing, the approximation is working well, so lambda is decreased. The parameter upgrade vector ( $u$ ) is calculated according to the following equation [58,60]:

$$u = (J^T \cdot W \cdot J + \lambda \cdot I)^{-1} \cdot J^T \cdot W \cdot r \quad (1)$$

where

$J$ —a Jacobean matrix, each element is a derivative of one particular observation for each parameter;

$W$ —a square diagonal matrix with squared weights attached to each observation;

$\lambda$ —Marquardt lambda or damping factor, which is adjusted during each iteration;

$r$ —vector of residuals which is minimized during calibration;

$I$ —an identity matrix; and

$T$ —the matrix transpose operator.

As it is apparent from Equation (1), the GLM algorithm requires the computation of the Jacobian matrix and matrix inversion at each iteration step as many times as there are adjustable parameters. Matrix inversion is a significant disadvantage. When the matrix is singular, GLM still performs due to the Marquardt lambda. A matrix often becomes singular when columns (or rows) are linearly dependent [61]. An enhanced version of GLM is implemented in PEST\_HP to improve computing performance. Each lambda defines a direction in parameter space; lambdas are selected using an enhanced version of the algorithm employed by the traditional PEST [62].

The GLM algorithm is a standard method for solving nonlinear least-squares problems. The method needs to adjust the parameterized functions to the set of measured data by reducing the sum of the

squares of the differences between the measured and computational values [44,60,63]. The result of this procedure is not only an optimal set of parameters but also a detailed analysis of the sensitivity of the parameters [11], which tells us what the impact and importance of each parameter are on the modeling itself [64]. Thus, several sets of parameter values that meet the criteria of the objective function with the same degree of adaptation can be obtained [30].

## 2.2. Singular Value Decomposition

When solving systems of linear equations, in practice, we may encounter very sensitive systems. The singular value decomposition (SVD) is a tool that helps us understand why a particular linear equation system is very sensitive. SVD is a very powerful tool in numerical linear algebra. With the help of SVD, we can get practical solutions to some very sensitive systems [45]. When a large number of parameters are added to a model, some parameters can be expected to be insensitive and others to be highly correlated with other parameters. As a result, even though a parameter may be estimable, it does not mean that it is estimable. What is needed is an intelligent calibration tool; one that detects what can and cannot be inferred from the calibration dataset. Then we can estimate what can be estimated and leave out what cannot be estimated. This is done automatically, without users' intervention. More about the theory of singular value decomposition (SVD) and the tool is described [56,65].

## 2.3. Tikhonov Regularization

The process by which a useful solution is obtained for a very sensitive system is called regularization [45]. Tikhonov regularization was first presented by Tikhonov and Arsenin as a solution finding procedure for ill-posed problems, which is often the case in models with a large number of parameters [43,66]. Using more parameters than can be constrained uniquely by observations results in the formulation of an ill-posed inverse problem. The numerical solution of that problem must include the use of one or more regularization mechanisms to stabilize the numerical solution process and identify a unique solution. This topic is covered in-depth in articles [67,68]. Boundary conditions constrain the inverse process calculations, and the experience of the modeler is required for their determination.

## 2.4. CMA-ES

The covariance matrix adaptation evolution strategy (CMA-ES) is a global optimization scheme. The CMA-ES is a stochastic or randomized method for the parameter calibration of nonlinear, nonconvex functions. CMA-ES is a powerful and robust global optimization method [38–40]. CMA-ES does not require derivatives of model outputs to adjustable parameters for the implementation of its optimization algorithm. Thus, it can be employed where model outputs show “numerical granularity” due to model numerical solution instability, or where the model is highly nonlinear and/or the objective function surface shows local minima at various scales [69].

## 2.5. GAP

Genetic algorithm and Powell (GAP) optimization is another approach for model calibration. The GAP is a stochastic design method and works evolutionarily by selecting and recombining high-performing parameter sets with each other. The GAP algorithm consists of two steps [18,54]. First, optimized parameter sets are generated by an evolutionary mechanism of selection and recombination of a set of initial, randomly selected parameter sets (again within user-defined parameter boundaries). During the second step, parameter sets are fine-tuned using Powell's quadratically convergent method, as described in [70].

## 2.6. Hydrological Model and Data

The hydrological model HBV in the version of HBV-light software [71] was selected for the study because of the possibility of its direct connection with external calibration programs. The HBV model is a widely used conceptual precipitation–runoff model that simulates catchment runoff based on precipitation, temperature, and long-term potential evaporation at a daily time step [72]. The HBV model was developed in the early 1970s to assist hydropower operations by the Hydrology Department of the Swedish Meteorological and Hydrological Institute [31]. The model is partially distributed since it allows the river catchment to be divided into smaller subcatchment units. Each subcatchment can be further subdivided into smaller areas based on land use and altitude. The model includes computational procedures that describe hydrological processes: snow accumulation and melting, evapotranspiration assessment and soil moisture calculation, subsoil runoff, and water flow transformation in a riverbed [73].

A version of the HBV-light model developed at Uppsala University was used and further refined at various universities around the world. The latest version of HBV-light was coded in the VB6 programming language in VB.NET [54]. In principle, the HBV-light routines correspond to the HBV model. However, some simplifications were made. Additional information on the HBV model is described in the articles [31,54,73].

An analysis based on a synthetic model was selected for the study, where the calibration procedure is performed based on the calculated catchment runoffs and taking into account known parameters. The process for creating the ideal model is relatively simple. The calibrated model is calculated once with known parameters. Then, the observed flow measurement ( $Q_{obs}$ ) is replaced by the simulated flow ( $Q_{sim}$ ) results. The model knows the simulated flow at known parameters, which is equal to the observed flow. The synthetic model eliminates measurement noise, conceptual noise, and model design noise from the calibration process. Therefore, only the calibration noise can be analyzed in this way. When calibrating such a model, a perfect match of the measured discharges with the simulated discharges can be expected [18,47]. However, when calibrated under ideal conditions, it is often not possible to re-establish (history match or to restore) the unique values of all parameters. In the synthetic model, deviations from perfect parameter matching are due to the noise generated by a particular calibration procedure and some nonunique parameters. This problem can be at least partially solved by appropriately parameterizing the relevant model equations [74].

For the calibration of parameters, the GLM calibration method, together with SVD and Tikhonov regularization, was used. Also, other calibration procedures, namely CMA–ES and GAP, were applied, expecting to result in a perfect goodness-of-fit measure as well. We selected these two additional procedures because they are already included in the software used for this study. CMA–ES is included in the PEST tool package, whereas GAP is integrated into the HBV-light program for calibration purposes. The GAP algorithm produces reasonably good results in finding a global solution to the observed function [18].

The existing hydrological test case included in the HBV-light model was selected for the analysis. For the test, Exercise 1 “HBV-Land” catchment case [54,55] was used. This hydrological test model needs 16 parameters to calibrate (Table 1). The model has more than enough parameters to test. The selected time step is one day in the period for ten years. The measured data on precipitation, temperature, and discharges are available for the period from 1 January 1981 to 31 December 1991. More details about HBV-Land catchment and the overall scenario of how to prepare and run a model are described in the article [54,55] in section A1 Exercise 1.

**Table 1.** List of parameters with their units and description.

Parameter Name	Unit <sup>1</sup>	Description
PERC	mm/ $\Delta t$	maximal flow from the upper to the lower box
UZL	mm	threshold parameter
K0	1/ $\Delta t$	storage (or recession) coefficient 0
K1	1/ $\Delta t$	storage (or recession) coefficient 1
K2	1/ $\Delta t$	storage (or recession) coefficient 2
MAXBAS	$\Delta t$	routing, length of the triangular weighting function
CET	1/ $^{\circ}\text{C}$	potential evaporation correction factor
TT	$^{\circ}\text{C}$	threshold temperature
CFMAX	mm/ $\Delta t$	degree- $\Delta t$ factor
SP	-	seasonal variability in degree- $\Delta t$ factor
SFCF	-	snowfall correction factor
CFR	-	refreezing coefficient
CWH	-	water-holding capacity
FC	mm	maximum soil moisture storage
LP	-	the threshold for reduction of evaporation (SM/FC)
BETA	-	a parameter that determines the relative contribution to runoff from rain or snowmelt

<sup>1</sup>  $\Delta t$  in hours or days, depending on the timescale of the model.

Since one of the main aims of this study is to analyze the calibration noise, we chose a synthetic model of a test example of the freely available HBV-Land included in the HBV-light. Moreover, with 16 parameters, we can analyze in more detail the impact of all the calibration procedures presented in the paper (i.e., GLM, CMA-ES, and GAP) to the known parameter values. If the analyses had been made on a model that had been calibrated with real measured output data, it would have been difficult to distinguish between the different types of noise that impede the result [26,30,42,75]. Moreover, another objective in this article is to test how to set up calibration parameters in a PEST control file to get satisfying results. A calibration failure in a primary case would put to question the continuation with a more complex model.

### 2.7. Model Calibration and Validation

Firstly, the GLM calibration method was performed. The procedure is schematically presented in Figure 1. The hydrological model with input data for the period from 1 January 1981 to 31 December 1991 and with known model parameters shown in the second column in Table 2 was calculated. For initial conditions, the model is expected to “warm-up” from 1 January 1981 to 30 October 1981. For the model, a computed value for the period from 1 October 1981 to 31 December 1991 was assumed. The calibration procedure was performed by using PEST [57] and PEST\_HP programs [76] running parallel on twelve logical processors on a single laptop PC. The procedure is step-by-step described in the Supplementary Materials. The calibration was repeated using the CMA-ES procedure and with the same control files settings as GLM (Figure 1).

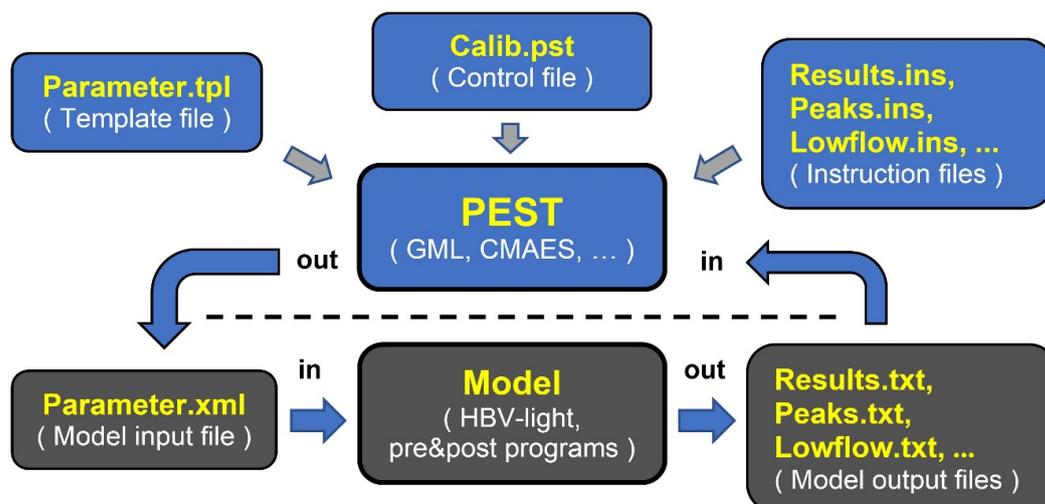


Figure 1. Calibration process workflow using the Parameter Estimation Tool—PEST.

Table 2. The values of initial parameters and boundary conditions.

Parameter Name	Perfect-Fit Value	Initial Value	Low Boundary	High Boundary
PERC	0.7	0.5	0.1	2
UZZ	20	30	5	50
K0	0.2	0.3	0.1	0.9
K1	0.08	0.1	0.01	0.2
K2	0.03	0.05	0.00005	0.1
MAXBAS	2.5	2	1	5
CET	0.1	0.2	0.01	0.3
TT	−1	1	−2	1
CFMAX	5	3	0.5	6
SP	1	0.9	0.01	1
SFCF	0.8	0.9	0.01	1
CFR	0.05	0.045	0.04	0.06
CWH	0.1	0.05	0.01	0.2
FC	250	200	50	500
LP	0.7	0.5	0.3	1
BETA	3	2	1	5

PEST is a suite of different programs for preparation, analysis, and control of calibration procedures and requires some knowledge and experience of this system. PEST programs require three types of support files (Figure 1): (1) parameter template files—one for each model input file that contains the parameters of variables that control the calibration; (2) instruction files—one file for each model output file from which the PEST must read the observed result numbers; and (3) one control file that provides the PEST with control parameters. The PEST control file includes the names of the relevant model input and output files, the operational mode like estimation or regularization, values of control variables, values of initial parameters, boundary ranges for each parameter, observed measurement values and weights. PEST runs Model, which executes the hydrological model and the necessary pre- and post-programs. The process is repeated until the criteria specified in the PEST control file are met. For more detailed procedures, see PEST instructions [62,69].

The same HBV-Land model, as for the GLM calibration method, was used to calibrate the HBV-light model by the GAP method. With the GAP evolutionary method, the procedure was repeated 100 times to run the model 5000 times, and 1000 times additionally with Powell local optimization [18,70]. When running the GAP procedure 100 times, 100 different parameter data sets are computed, and then the appropriate set regarding the selected goodness-of-fit criteria in Table 3 is used. Moreover, the GAP procedure was repeated to run 100 times, but with only 13 parameters (GAP<sub>13</sub>). In this way,

it was tested if the parameter values are closer to the initial ones than by using all 16 parameters ( $GAP_{16}$ ), where SP, CFR, and CWH values listed in Table 1, were fixed to the perfect-fit values (Table 2). A similar fixing of some of the parameters to avoid over-parameterization was done with CFR and CWH described in the article [18], where K0 and UZL parameters were not used in the calibration process with GAP.

For the initial values of the parameters, the values listed in the third column in Table 2 are taken into account. All calibration procedures, i.e., GLM, CMA-ES, and GAP, are run with the same initial and boundary conditions shown in Table 2.

During the calibration process, composite sensitivities were calculated and normalized for all observations. Normalized sensitivity shows how the uncertainty in the output of a model can be apportioned to different input parameters.

Model validation was performed using measurements of precipitation and temperature for the period from 1 January 2001 to 31 December 2011. Similar to the calibration procedure [54,55], the first nine months of the calculation are used to warm-up the model. Discharge data are simulated with known parameters from the calibration for the period from 1 October 2001 to 31 December 2011. This validation model is the same as in the calibration process, but with different input of precipitation and temperature data. Evapotranspiration data are the same as in the calibration period.

The efficiency of a model is commonly evaluated by the Nash–Sutcliffe efficiency (NSE) coefficient, which has been used as a standard for decades in testing calibration and model validation [29]. The quality of the calibrated model is shown in the validation process. The results of the computations are compared with sample data that were not used in the calibration. A successful calibration model is expected to achieve a similar NSE coefficient to that in the calibration when validating.

In addition to the Nash–Sutcliffe efficiency, other measures of the model calibration performance were taken into account in the analysis, namely the coefficient of determination, Kling–Gupta efficiency [77], efficiency for log (Q), flow-weighted efficiency, mean difference, efficiency for peak flows, efficiency at low flows, low-flow difference, and volume error (Table 3). Equations of efficiency for peak flow, model efficiency ( $R_{eff}$ ), and efficiency for low flow are the same as NSE coefficients, differing just in different data series periods. For a peak flow, a data point with the highest observed discharge ( $Q_{obs}$ ) value that is at least three times the average  $Q_{obs}$  within a time window of 15 days is used. Nevertheless, the most fundamental hydrological characteristic is the low flow, which is a seasonal phenomenon and an integral part of every stream [78]. Low flows were identified as challenging to calibrate in several studies [79,80]. For low flows, the arbitrary upper bound is given by the mean annual runoff, which is a mean value of the available flow time series of the annual flow.  $Q_{70}$  flows within the range of 70%-time exceedance threshold were used as low flows in this study. The same threshold was selected in the low-flow study in a heterogeneous catchment in Slovenia [81]. For the HBV-Land catchment model, based on the calibration period data from 1 October 1981 to 31 December 1991, the calculated low-flow threshold  $Q_{70}$  is 0.39 mm.

**Table 3.** Coefficients for calibration quality assessments, goodness-of-fit measures according to [82].

Benchmark Description	Definition <sup>1</sup>	Value for Perfect-Fit	Range
Coefficient of determination (R <sup>2</sup> )	$\frac{(\sum (Q_{obs} - \overline{Q_{obs}})(Q_{sim} - \overline{Q_{sim}}))^2}{\sum (Q_{obs} - \overline{Q_{obs}})^2 \sum (Q_{sim} - \overline{Q_{sim}})^2}$	1	−∞ to 1
Nash–Sutcliffe efficiency (NSE)	$1 - \frac{\sum (Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - \overline{Q_{obs}})^2}$	1	−∞ to 1
Efficiency for log(Q)	$1 - \frac{\sum (\ln Q_{obs} - \ln Q_{sim})^2}{\sum (\ln Q_{obs} - \ln \overline{Q_{obs}})^2}$	1	−∞ to 1
Flow-weighted efficiency	$1 - \frac{\sum (Q_{obs}/\overline{Q_{obs}})(Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - \overline{Q_{obs}})^2}$	1	−∞ to 1
Mean difference	$\frac{\sum (Q_{obs} - Q_{sim})}{n} 365$	0	−∞ to ∞
Efficiency for peak flows	$1 - \frac{\sum (peak Q_{obs} - peak Q_{sim})^2}{\sum (peak Q_{obs} - peak \overline{Q_{obs}})^2}$	1	−∞ to 1
Volume error	$1 - \frac{1}{n} \sum  Q_{obs} - Q_{sim} $	1	−∞ to 1
MARE measure	$1 - \frac{1}{n} \sum \frac{ Q_{obs} - Q_{sim} }{Q_{obs}}$	1	−∞ to 1
Lindstrom measure	$NSE - 0.1 \frac{ \sum (Q_{obs} - Q_{sim}) }{\sum (Q_{obs})}$	1	−∞ to 1
Spearman rank	$\frac{\sum (R_{obs} - \overline{R_{obs}})(S_{sim} - \overline{S_{sim}})}{\sqrt{\sum (R_{obs} - \overline{R_{obs}})^2} \sqrt{\sum (S_{sim} - \overline{S_{sim}})^2}}$	1	−∞ to 1
The efficiency of low flows	$1 - \frac{\sum (low Q_{obs} - low Q_{sim})^2}{\sum (low Q_{obs} - low \overline{Q_{obs}})^2}$	1	−∞ to 1
Low-flow difference	$\sum (low Q_{obs} - low Q_{sim})$	0	−∞ to ∞

<sup>1</sup>  $Q_{obs}$  is the measured (observed) flow,  $Q_{sim}$  is the modeled (simulated) flow,  $n$  is the number of time steps, and  $R_{obs}$  and  $S_{sim}$  are the ranks of  $Q_{obs}$  and  $Q_{sim}$ .

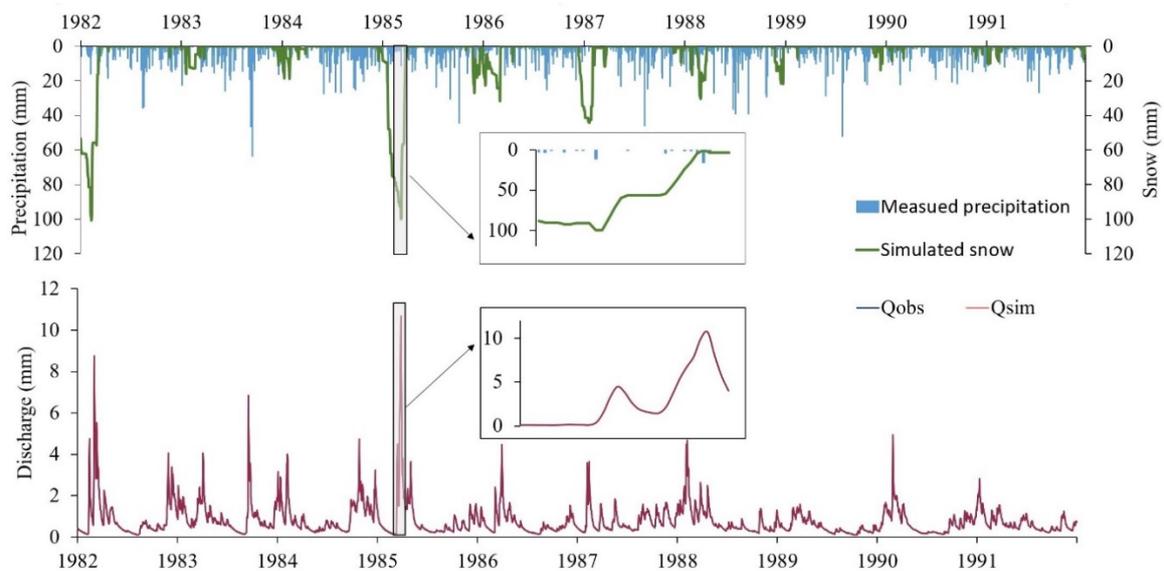
Usually, the coefficient of efficiency ( $R_{eff}$ ) is used in the HBV model.  $R_{eff}$  is the Nash–Sutcliffe efficiency coefficient (NSE) of a whole calculated period and can range from  $-\infty$  to 1. When  $NSE = 1$ , then we get a “perfect-fit value” and it means that  $Q_{sim}(t) = Q_{obs}(t)$ . When  $NSE = 0$ , the simulation is as good or as poor as the constant-value prediction. On the other hand,  $NSE < 0$  means an inferior fit. As can be seen from the equations in Table 3, all discharges are time-dependent. The NSE coefficient is sensitive to extreme values. Model relevance can be objectively accepted or rejected based on the NSE greater than the threshold selected by the user. The acceptable values depend on the use of the model and the practice of modeling. Results of calculations presented in this paper showed a good approximation of the perfect-fit values, i.e.,  $NSE = 1$ .

Moreover, the parameter’s relative deviation (%) was calculated by comparing the simulated parameter ( $P_{sim}$ ) to perfect-fit parameter ( $P_{PF}$ ) to standardize the various metrics and allow for a comparison of the parameters (Equation (2)).

$$\text{Parameter relative deviation (\%)} = (P_{PF} - P_{sim})/P_{PF} \times 100 \tag{2}$$

### 3. Results and Discussion

Calibration results with the GLM calibration procedure are graphically presented in Figure 2. Results of efficiency measures in the synthetic model with perfect parameter values are given in the second column of Table 4 with the label “Perfect-fit value.” The third column in Table 4 shows the results of parameters from the calibration using the GLM algorithm with the use of SVD and Tikhonov regularization. For the GLM procedure, the run time was less than 1 min. Despite the same calculation data, slightly different results were obtained, which can be attributed to the computation noise. This noise was most likely caused by rounding the  $Q_{sim}$  results in the calculation.



**Figure 2.** Observed and simulated discharges with the Gauss–Levenberg–Marquardt method (GLM) calibration procedure for the calibration period from 1 January 1982 to 31 December 1991.

**Table 4.** Calibration results <sup>1</sup> from history matching of the synthetic model over ten years, i.e., 1 October 1981 to 31 December 1991.

Model Efficiency Benchmark <sup>1</sup>	Perfect-Fit Value	GLM	CMA–ES	GAP <sub>16</sub>	GAP <sub>13</sub>
Coefficient of determination R2	1	1	0.999999	0.999631	0.99999
Model efficiency R <sub>eff</sub> (NSE)	1	1	0.999999	0.999628	0.999989
Kling–Gupta efficiency KGE	0.999990	0.999990	0.999902	0.998184	0.999003
Kling–Gupta efficiency “nonparametric”	0.999844	0.999840	0.999774	0.997672	0.999156
Efficiency for log(Q)	1	1	0.999996	0.998712	0.999946
Flow-weighted efficiency	1	1	0.999999	0.999777	0.999994
Mean difference	0.002825	0.002860	0.023888	-0.53402	0.12956
Efficiency for peak flows	1	1	0.999997	0.998706	0.999978
Volume error	0.999990	0.999990	0.999919	0.998193	0.999562
MARE measure	0.999453	0.999450	0.999348	0.990692	0.996568
Lindstrom measure	1	1	0.999991	0.999447	0.999946
Spearman rank	1	1	0.999999	0.999688	0.999987
Efficiency of low flows	1	1	0.999969	0.983838	0.999292
Low-flow difference	0	0	0.044	-1.307	-1.063

<sup>1</sup> Results are rounded to the sixth decimal.

The NSE of the computation is quite high, as shown in the third column in Table 4. Deviations merely appeared after the tenth decimal place in the NSE (NSE perfect-fit value = 0.9999998607; NSE GLM = 0.9999998605) and after the fifth decimal place in the mean differences benchmark, which can be almost equated with the perfect-fit value. The performance of calibration with CMA–ES, GAP<sub>13</sub>, and GAP<sub>16</sub> is shown in columns 4–6 of Table 4. The overall comparison of NSE of all four procedures applied in this study is shown in Figure 3. Based on NSE, one can notice an excellent performance also of CMA–ES and GAP<sub>13</sub>, while GAP<sub>16</sub> reached the lowest NSE.

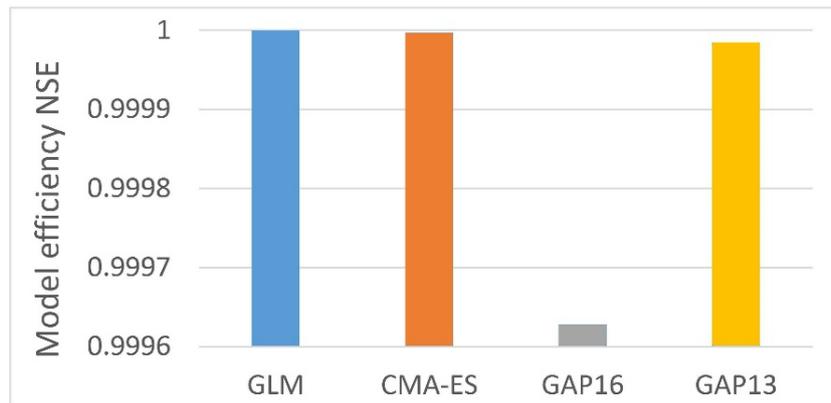


Figure 3. Presentation of NSE results for the calibration period.

PEST\_HP calculates parameter sensitivity during the calibration process itself. The normalized sensitivity in the second column of Table 5 shows how the uncertainty in the output of a model can be distributed to different parameters. The TT (temperature threshold) parameter is the most sensitive parameter with normalized sensitivity 1. In contrast, the least sensitive is CFR (refreezing coefficient) and CWH (water-holding capacity) parameters with normalized sensitivities of 0. The results of the calculated parameters with GLM, CMA-ES, and GAP are shown in columns 3–6 of Table 5. The results calibrated by the GLM procedure in the third column of Table 5 give a value close to the perfect-fit parameter value of Table 2. The differences occur only from the fourth to the fifth decimal place as a result of the calibration noise shown in column 3 of Table 5. The parameter values calibrated by the GAP procedure give relatively good results (GAP<sub>16</sub>, NSE = 0.9996) compared to the existing hydrological practice. However, they are significantly worse compared to GLM, as shown in Tables 5 and 6.

Table 5. The values of the individual parameters determined in the calibration procedure.

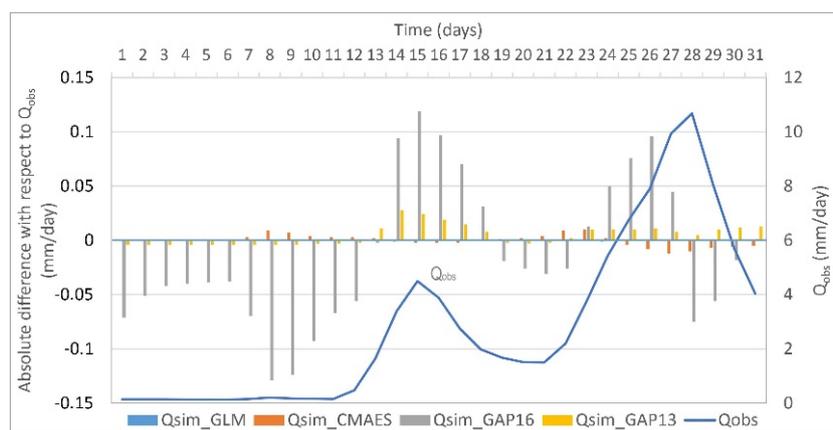
Parameter Name	Normalized Sensitivity	GLM	CMA-ES	GAP <sub>16</sub>	GAP <sub>13</sub>
PERC	0.03	0.700005	0.7000623	0.697142508	0.693526262
UZL	0.07	20.0001	19.99882	19.95660857	20.11814985
K0	0.03	0.20001	0.1998579	0.204047194	0.20075344
K1	0.04	0.08	0.08000101	0.079860003	0.0799286
K2	0.01	0.03	0.0300044	0.030148466	0.029351392
MAXBAS	0.07	2.49999	2.499148	2.512553604	2.500201514
CET	0.08	0.100004	0.100003	0.102796635	0.100969107
TT	1.00	−0.99998	−1.000009	−0.99999999	−0.99999592
CFMAX	0.02	5.00016	5.034117	5.346262187	4.984995411
SP	0.13	0.999978	0.9910489	0.879212318	-
SFCF	0.23	0.799992	0.7997364	0.813747468	0.797604941
CFR	0.00	0.050009	0.05129717	0.044801671	-
CWH	0.00	0.100002	0.1000745	0.064851987	-
FC	0.11	250	250.0022	244.8327906	249.3696183
LP	0.12	0.700005	0.7000338	0.677129184	0.700118925
BETA	0.06	3.000014	3.000171	2.871328415	3.005579383

**Table 6.** Parameter relative deviation to “perfect fit” in the calibration procedure.

Parameter Name	GLM %	CMA-ES %	GAP <sub>16</sub> %	GAP <sub>13</sub> %
PERC	0	−0.01	0.41	0.92
UZL	0	0.01	0.22	−0.59
K0	0	0.07	−2.02	−0.38
K1	0	0	0.17	0.09
K2	0	−0.01	−0.49	2.16
MAXBAS	0	0.03	−0.5	−0.01
CET	0	0	−2.8	−0.97
TT	0	0	0	0
CFMAX	0	−0.68	−6.93	0.3
SP	0	0.9	12.08	-
SFCF	0	0.03	−1.72	0.3
CFR	−0.02	−2.59	10.4	-
CWH	0	−0.07	35.15	-
FC	0	0	2.07	0.25
LP	0	0	3.27	−0.02
BETA	0	−0.01	4.29	−0.19

The advantage of using the GLM with SVD and Tikhonov regularization in the calibration process is evident by comparing the values of individual relative deviations concerning the “perfect-fit” of the parameters in Table 6.

In the regularization process (GLM), almost a perfect match between the true and the calibrated parameters is achieved, deviating by only 0.02% of CFR (refreezing coefficient) shown in the second column in Table 6. This small deviation (noise) is probably a consequence of the computation. Differences occur after the third decimal place. They can be explained by the equation for calculating refreezing meltwater (RM) defined as  $RM = CFR \cdot CFMAX \cdot (TT - T(t))$  and containing three parameters and one variable [18,54,72,83]. If a parameter such as CFR has independent relationships with other parameters, it cannot be adequately calculated due to over-parameterisation. It is therefore advisable to determine the value of such a parameter based on the expertise of the modeler. As the catchment has no altitude zones, the occurrence of refreezing meltwater was also relatively rare in the observed period. The results of the calibration procedure with GLM are graphically presented in Figures 4 and 5.



**Figure 4.** The absolute difference to  $Q_{obs}$  for different calibration procedures in the calibration period from 1 March 1985 to 31 March 1985.

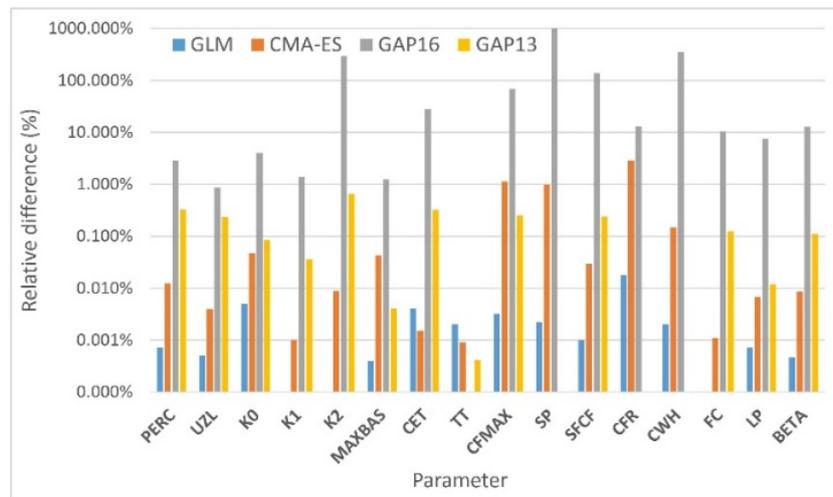


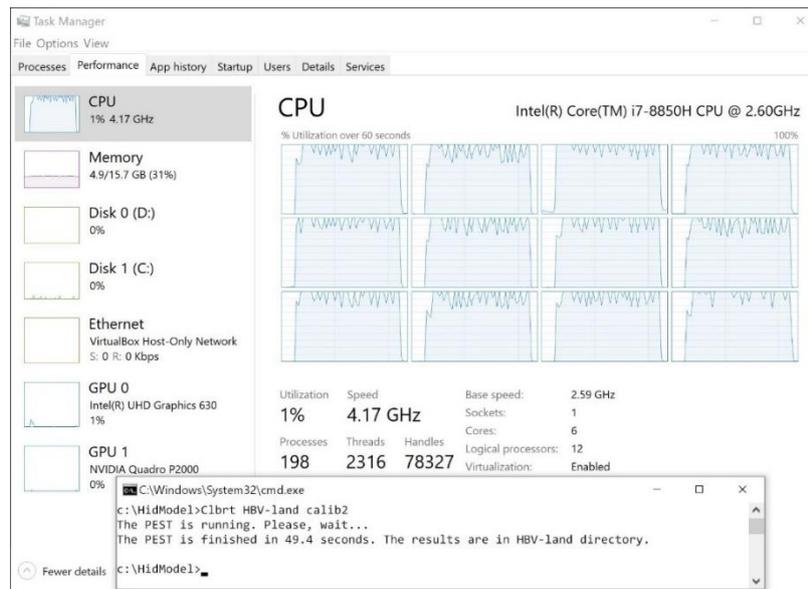
Figure 5. Graphical presentation of calibrated parameters with different optimizers.

On the other hand, parameter values obtained by calibration using the GAP algorithm differ significantly from the true values used for the calculation of the HBV-Land model (Figures 4 and 5, and Table 6). In the case of GAP<sub>16</sub>, 10 of 16 parameters deviate by more than 1%; SP, CFR, and CWH parameters deviate even by 12%, 10%, and 35%, respectively. Moreover, the total parameter deviation obtained by the GAP<sub>16</sub> procedure is 82%. CMA-ES performed better than GAP but worse than GLM. The largest deviation (2.6%) in CMA-ES is again observed for the CFR parameter. The second-largest deviation (0.7%) is found for the CFMAX parameter. The total deviation of parameters in case of CMA-ES is 6%.

When the number of parameters in a calibration set was reduced from 16 to 13, a more efficient GAP calibration with the total deviation of parameters of 6% resulted. One hundred optimized parameter sets resulted in a maximum, minimum, and median NSE equal to 0.999989, 0.980133, and 0.994431, respectively. For comparing GAP<sub>16</sub> with other calibration methods, the parameter set obtained with maximum NSE was selected.

Many researchers found that low-flow calibration is quite problematic [79,80]. In this study, the only calibration using the GLM procedure gave a perfect fit in low flows (Table 4) without using the additional criteria to calibrate low-flow data series.

CMA-ES parameter relative deviation to history matching is close to GLM, as shown in the second column of Table 6. The calibration time can sometimes be significant. The run time depends on computer power, the operating system, and calibration process termination criteria. We need to look at this metric relatively and see how much one method is faster or slower than the other. For the selected case of the hydrological model calibration, the calculation using the parallel computing according to the GLM procedure takes 50 s (Figure 6), the CMA-ES parallel calculation takes one hour, and one GAP calculation takes two minutes with 6000 model runs. To achieve the results shown in Table 6, one hundred repetitions of GAP calculations take about three hours and ten minutes.



**Figure 6.** CPU utilization when HBV-Land is calibrated with parallel PEST.

The GLM procedure calculated the hydrological model 533 times with 14 iterations, and the CMA-ES ran the model 12,937 times with 925 iterations. For one hundred repetitions with the GAP procedure, 603,150 model runs of the model are required. Better efficiency of the GLM procedure regarding the number of required model calculations in comparison with the methods of evolutionary CMA-ES and GAP is hence more than evident. Moreover, the GLM calibration process in all cases calculated the same NSE value and values of calibrated parameters remained the same, regardless of the number of repetitions of the calibration process. The same repeating results can be attributed to the nonstochastic behavior in GLM calibration procedures [84], in contrast to stochastic evolutionary procedures (CMA-ES and GAP in this study). The minimum error variance in the objective function was obtained with the GLM procedure in all cases.

The validation results were similar to those of the calibration but slightly worse for all procedures. However, especially for models calibrated using GAP, the differences are relatively more considerable. Validation results with the GLM calibration procedure are graphically presented in Figure 7. For the model based on the GLM calibration, the NSE for the calibration and validation period has the same value  $NSE = 1$  (Table 7). For the model based on the GAP calibration parameters, NSE decreases from 0.9996 for the calibration period to 0.9983 for the validation period (Table 7).

Brilly et al. [85], as part of the broader study, reported about using the proposed combined procedure, implemented in the PEST software and on a real catchment, i.e., the Savinja River catchment in Slovenia. The average NSE value for the model calibration of 21 subcatchments was 0.85, suggesting that the proposed methodology has great potential also for modeling (multiple) real river catchments. However, this is the first paper testing the use of the combined GLM procedure on a synthetic rainfall-runoff model, where all other types of noise, except calibration noise, are excluded from the model. A comparison of results of all four calibration approaches used in this study shows that calibration noise was excluded only with the combined GLM method.

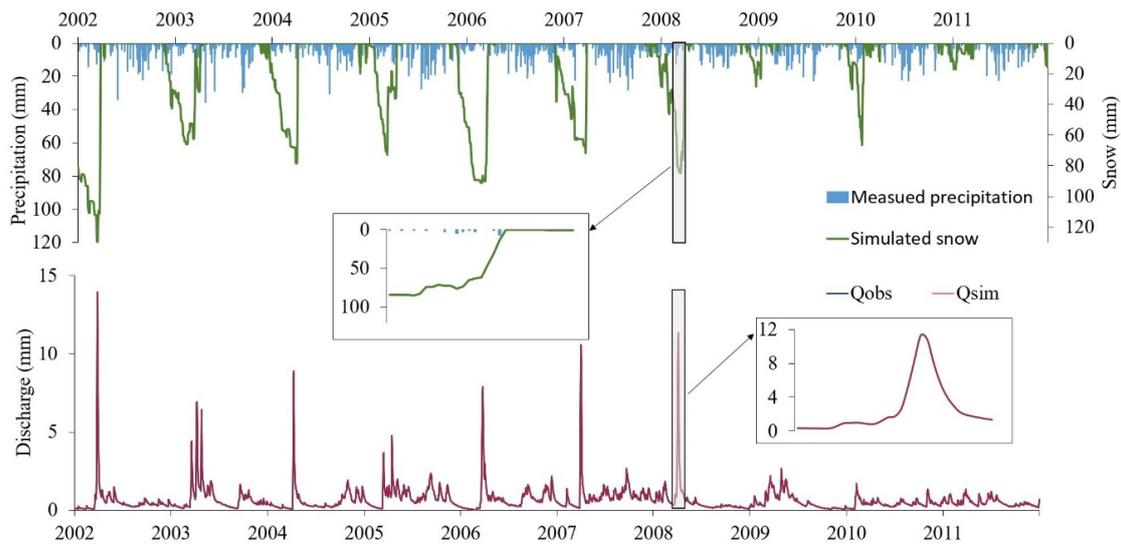


Figure 7. Validation period 1 January 2002 to 31 December 2011, validated with GLM.

Table 7. Validation results <sup>1</sup> of the HBV-Land model for the period 1 October 2001 to 31 December 2011.

Model Efficiency Benchmark	Perfect-Fit Value	GLM	CMA-ES	GAP <sub>16</sub>	GAP <sub>13</sub>
Coefficient of determination R2	1	1	0.999995	0.998259	0.999989
Model efficiency NSE	1	1	0.999994	0.998092	0.999987
Kling-Gupta efficiency KGE	0.999987	0.99998	0.99924	0.984518	0.998326
Kling-Gupta efficiency “non-parametric” NPE	0.999811	0.99981	0.999562	0.987471	0.998312
Efficiency for log(Q) LogR <sub>eff</sub>	1	1	0.999992	0.993776	0.999872
Flow-weighted efficiency FlowWeightedNSE	1	1	0.999995	0.998288	0.999996
Mean difference	-0.001665	-0.0035	0.029267	-2.79455	-0.31087
Efficiency for peak flows	1	1	0.999982	0.995386	0.999985
Volume error	0.999993	0.99999	0.999879	0.988476	0.998718
MARE measure	0.999205	0.99921	0.998851	0.970793	0.994208
Lindstrom measure	1	1	0.999982	0.99694	0.999859
Spearman rank	1	1	0.999996	0.998679	0.999981
Efficiency for low flows	1	1	0.999977	0.980824	0.999784
Low-flow difference	0.001	-0.007	0.232	-11.999	-2.27

<sup>1</sup> Results are rounded to the sixth decimal.

#### 4. Conclusions

In this paper, the enhanced Gauss–Levenberg–Marquardt (GLM) procedure in combination with the singular value decomposition (SVD) and Tikhonov regularization is presented and applied to calibrate the conceptual hydrological model. The procedure was evaluated and compared with two other global calibration techniques (GAP and CMA–ES). Based on the results of this study, the following conclusions can be drawn regarding the presented combined use of GLM, SVD, and Tikhonov regularization:

- The combined GLM procedure enabled an almost perfect match between true and calculated parameters. Only one parameter out of 16 showed a deviation of only 0.02%.
- By knowing the true model parameters, various calibration parameter settings can be tested. With the appropriate parameter settings of calibration, the calibration noise was reduced, and the calibration efficiency was improved.
- Only with the GLM procedure, the precise values of parameters on the HBV-Land model for both peak and low flows were obtained. In other words, only the combined GLM process was able to eliminate the calibration noise in the case of simulations based on the synthetic model runoff.
- The run time of the combined GLM procedure was significantly shorter than those of the compared procedures. For the same case study, the combined GLM calibration procedure was approximately 60 and 200-times shorter than of CMA–ES and GAP, respectively.

- The GLM calibration process in all cases calculated the same NSE value, and values of calibrated parameters remained the same, regardless of the number of repetitions of the calibration process. The same repeating results can be attributed to the nonstochastic behavior in GLM calibration procedures [84], in contrast to stochastic evolutionary procedures (CMA-ES and GAP in this study). The minimum error variance in the objective function was obtained with the GLM procedure in all cases.
- The noise in the calculation results with the GLM method was practically the same in either calibrating or validating the procedure. In the modeling, therefore, only computational noise remained in the results.
- The result of calibration with the GLM procedure, which would be stuck at the local minimum, was not detected in this study. The GLM procedure has proven to be very useful in solving linear inverse problems. By introducing Tikhonov regularization into an inverse solution, we successfully calculated unique parameters. With successive iterations, the GLM method achieved model calibration, regardless of the nonlinear relationship between model outputs and model parameters.
- The calibration of the synthetic model gave an insight into the noise generated and the deficiencies in the design of the calibration process. However, the GLM calibration process itself requires expertise in hydrological modeling as well as in the calibration process. The lack of expertise is also the reason that the GLM process has not yet been widely implemented in practice to calibrate distributed hydrological models. For this reason, a Supplementary Materials section was prepared to bring the procedure closer to potential users.

Results of this study suggest that the proposed procedure (detailed explanation in the Supplementary Materials—A Quickstart Guide to HBV-Light Hydrological Model Calibration Using PEST) has a great potential to address many open challenges, such as calibration of multiple subcatchments at once instead of one to the other and from top to bottom.

**Supplementary Materials:** The following is available online at <http://www.mdpi.com/2076-3417/10/11/3841/s1>, A Quickstart Guide to HBV-Light Hydrological Model Calibration Using PEST.

**Author Contributions:** Conceptualization, A.V. and M.B.; methodology, A.V.; software, A.V.; validation, A.V., M.B. and K.S.; formal analysis, A.V. and M.B.; investigation, A.V.; resources, A.V.; data curation, A.V.; writing—original draft preparation, A.V., M.B., K.S., and A.K.; writing—review and editing, A.V. and K.S.; visualization, A.V.; supervision, M.B. and A.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Slovenian Research Agency (ARRS), grant number P2-0180 (Water Science and Technology, and Geotechnical Engineering: Tools and Methods for Process Analyses and Simulations, and Development of Technologies).

**Acknowledgments:** The authors thank John Doherty for his valuable advice while working on this research. We appreciate the ARRS Slovenian Research Agency and the Slovenian UNESCO Commission for their financial support for the work on the assignment and presentations at scientific meetings. We appreciate the anonymous reviewers for their valuable comments and efforts to improve this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Singh, V.P.; Donald, K. *Frevert Watershed Models (Hardback)*; Taylor & Francis: Oxfordshire, UK; CRC Press: Boca Raton, FL, USA, 2005; ISBN 9780849336096.
2. Beven, K. *Rainfall-Runoff Modelling: The Primer*, 2nd ed.; Beven, K., Ed.; John Wiley & Sons, Ltd.: Boca Raton, FL, USA, 2010; pp. 157–175. ISBN 978-0-470-71459-1.
3. McCulloch, J.S.G. Hydrology in practice. *J. Hydrol.* **1994**, *160*, 141–142. [[CrossRef](#)]
4. Refsgaard, J.C.; Storm, B. *Construction, Calibration And Validation of Hydrological Models*; Springer: Dordrecht, The Netherlands, 1990; pp. 41–54. [[CrossRef](#)]
5. Reggiani, P.; Schellekens, J. Modelling of hydrological responses: The representative elementary watershed approach as an alternative blueprint for watershed modelling. *Hydrol. Process.* **2003**, *17*, 3785–3789. [[CrossRef](#)]

6. Westra, S.; Thyer, M.; Leonard, M.; Kavetski, D.; Lambert, M. A strategy for diagnosing and interpreting hydrological model nonstationarity. *Water Resour. Res.* **2014**, *50*, 5090–5113. [[CrossRef](#)]
7. Abbott, M.B.; Bathurst, J.C.; Cunge, J.A.; O’Connell, P.E.; Rasmussen, J. An introduction to the European Hydrological System—Systeme Hydrologique Europeen, “SHE”, 2: Structure of a physically-based, distributed modelling system. *J. Hydrol.* **1986**, *87*, 61–77. [[CrossRef](#)]
8. Peel, M.C.; Blöschl, G. Hydrological modelling in a changing world. *Prog. Phys. Geogr. Earth Environ.* **2011**, *35*, 249–261. [[CrossRef](#)]
9. Kavetski, D.; Kuczera, G.; Franks, S.W. Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts. *J. Hydrol.* **2006**, *320*, 173–186. [[CrossRef](#)]
10. Kuczera, G.; Renard, B.; Thyer, M.; Kavetski, D. Il n’y a pas de monstres hydrologiques, juste des modèles et des observations avec de grandes incertitudes! *Hydrol. Sci. J.* **2010**, *55*, 980–991. [[CrossRef](#)]
11. Doherty, J.; Johnston, J.M. No Title. *J. Am. Water Resour. Assoc.* **2003**, *39*, 251–265. [[CrossRef](#)]
12. Freeze, R.A.; Harlan, R.L. Blueprint for a physically-based, digitally-simulated hydrologic response model. *J. Hydrol.* **1969**, *9*, 237–258. [[CrossRef](#)]
13. Refsgaard, J.C.; Knudsen, J. Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.* **1996**, *32*, 2189–2202. [[CrossRef](#)]
14. Graham, D.N.; Butts, M.B. Flexible integrated watershed modeling with MIKE SHE. In *Watershed Models*; CRC Press: Boca Raton, FL, USA, 2005; pp. 245–271. ISBN 9781420037432.
15. Merz, R.; Parajka, J.; Blöschl, G. Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resour. Res.* **2011**, *47*. [[CrossRef](#)]
16. Andréassian, V.; Lerat, J.; Loumagne, C.; Mathevet, T.; Michel, C.; Oudin, L.; Perrin, C. What is really undermining hydrologic science today? *Hydrol. Process.* **2007**, *21*, 2819–2822. [[CrossRef](#)]
17. Beven, K. How to make advances in hydrological modelling. *Hydrol. Res.* **2019**, *50*, 1481–1494. [[CrossRef](#)]
18. Seibert, J. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrol. Earth Syst. Sci.* **2000**, *4*, 215–224. [[CrossRef](#)]
19. Chintalapudi, S.; Sharif, H.; Xie, H. Sensitivity of Distributed Hydrologic Simulations to Ground and Satellite Based Rainfall Products. *Water* **2014**, *6*, 1221–1245. [[CrossRef](#)]
20. Shedekar, V.S.; King, K.W.; Brown, L.C.; Fausey, N.R.; Heckel, M.; Harmel, R.D.; Reno, N. *Measurement Errors in Tipping Bucket Rain Gauges under Different Rainfall Intensities and their implication to Hydrologic Models*; American Society of Agricultural and Biological Engineers: St. Joseph, MI, USA, 2009.
21. Dymond, J.R.; Christian, R. Accuracy of discharge determined from a rating curve. *Hydrol. Sci. J.* **1982**, *27*, 493–504. [[CrossRef](#)]
22. Bonacci, O. The influence of errors in precipitation measurements on the accuracy of the evaporation measurements performed by a class A evaporation pan. *Theor. Appl. Climatol.* **1991**, *43*, 181–183. [[CrossRef](#)]
23. Sevruk, B. *Methods of Correction for Systematic Error in Point Precipitation Measurement for Operational Use*; Secretariat of the World Meteorological Organization: Geneva Switzerland, 1982; ISBN 9789263105899.
24. Westerberg, I.K.; Wagener, T.; Coxon, G.; McMillan, H.K.; Castellarin, A.; Montanari, A.; Freer, J. Uncertainty in hydrological signatures for gauged and ungauged catchments. *Water Resour. Res.* **2016**, *52*, 1847–1865. [[CrossRef](#)]
25. Kuczera, G.; Kavetski, D.; Franks, S.; Thyer, M. Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *J. Hydrol.* **2006**, *331*, 161–177. [[CrossRef](#)]
26. Doherty, J.; Welter, D. A short exploration of structural noise. *Water Resour. Res.* **2010**, *46*. [[CrossRef](#)]
27. Beven, K.; Binley, A. The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Process.* **1992**, *6*, 279–298. [[CrossRef](#)]
28. Shaw, E.M.; Beven, K.J.; Chappell, N.A.; Lamb, R. *Hydrology in Practice*, 4th ed.; CRC Press Hydrology in Practice: New York, NY, USA, 2017; ISBN 9781482265705. [[CrossRef](#)]
29. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
30. Kavetski, D.; Kuczera, G.; Franks, S.W. Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis. *J. Hydrol.* **2006**, *320*, 187–201. [[CrossRef](#)]
31. Bergström, S. Development and Application of a Conceptual Runoff Model for Scandinavian Catchments. *Smhi* **1976**, *7*, 134. [[CrossRef](#)]

32. Bergström, S. *The HBV Model—Its Structure and Applications*; Swedish Meteorological and Hydrological Institute: Norrköping, Sweden, 1992.
33. Zhang, Y.; Shao, Q.; Zhang, S.; Zhai, X.; She, D. Multi-metric calibration of hydrological model to capture overall flow regimes. *J. Hydrol.* **2016**, *539*, 525–538. [[CrossRef](#)]
34. Solomatine, D.P.; Shrestha, D.L. A novel method to estimate model uncertainty using machine learning techniques. *Water Resour. Res.* **2009**, *45*, 1–16. [[CrossRef](#)]
35. Freer, J.; Beven, K.; Ambroise, B. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resour. Res.* **1996**, *32*, 2161–2173. [[CrossRef](#)]
36. Shrestha, D.L.; Kayastha, N.; Solomatine, D.P. A novel approach to parameter uncertainty analysis of hydrological models using neural networks. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 1235–1248. [[CrossRef](#)]
37. Karahan, H.; Gurarslan, G.; Geem, Z.W. Parameter estimation of the nonlinear muskingum flood-routing model using a hybrid harmony search algorithm. *J. Hydrol. Eng.* **2013**, *18*, 352–360. [[CrossRef](#)]
38. Hansen, N.; Ostermeier, A. Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* **2001**, *9*, 159–195. [[CrossRef](#)]
39. Hansen, N.; Müller, S.D.; Koumoutsakos, P. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol. Comput.* **2003**, *11*, 1–18. [[CrossRef](#)] [[PubMed](#)]
40. Hansen, N.; Hansen, N. The CMA Evolution Strategy: A Comparing Review. *STUDFUZZ* **2006**, *192*, 75–102.
41. Ayvaz, M.T.; Gurarslan, G. A new partitioning approach for nonlinear Muskingum flood routing models with lateral flow contribution. *J. Hydrol.* **2017**, *553*, 142–159. [[CrossRef](#)]
42. Kavetski, D.; Qin, Y.; Kuczera, G. The Fast and the Robust: Trade-Offs Between Optimization Robustness and Cost in the Calibration of Environmental Models. *Water Resour. Res.* **2018**, *54*, 9432–9455. [[CrossRef](#)]
43. Willoughby, R.A. Solutions of Ill-Posed Problems (A. N. Tikhonov and V. Y. Arsenin). *SIAM Rev.* **1979**, *21*, 266–267. [[CrossRef](#)]
44. Doherty, J.; Skahill, B.E. An advanced regularization methodology for use in watershed model calibration. *J. Hydrol.* **2006**, *327*, 564–577. [[CrossRef](#)]
45. Plestenjak, B. *Razširjen Uvod v Numerične Metode*; DMFA—Založništvo: Ljubljana, Slovenija, 2015; ISBN 9789612122645.
46. Welter, D.E.; White, J.T.; Hunt, R.J.; Doherty, J.E. *Approaches in Highly Parameterized Inversion: PEST ++ Version 3, a Parameter ESTimation and Uncertainty Analysis Software Suite Optimized for Large Environmental Models*; U.S. Geological Survey: Reston, VA, USA, 2015. [[CrossRef](#)]
47. White, J.T. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. *Environ. Model. Softw.* **2018**, *109*, 191–201. [[CrossRef](#)]
48. Fioren, M.N.; Masterson, J.P.; Plant, N.G.; Gutierrez, B.T.; Thiel, E.R. Bridging groundwater models and decision support with a Bayesian network. *Water Resour. Res.* **2013**, *49*, 6459–6473. [[CrossRef](#)]
49. Harrison, M.T.; Roggero, P.P.; Zavattaro, L. Simple, efficient and robust techniques for automatic multi-objective function parameterisation: Case studies of local and global optimisation using APSIM. *Environ. Model. Softw.* **2019**, *117*, 109–133. [[CrossRef](#)]
50. Li, X.; Cao, J.; Du, D. Comparison of Levenberg-Marquardt method and path following interior point method for the solution of optimal power flow problem. *Int. J. Emerg. Electr. Power Syst.* **2013**, *13*. [[CrossRef](#)]
51. Bjarkason, E.K.; Maclaren, O.J.; Nicholson, R.; Yeh, A.; O'sullivan, M.J. Uncertainty Quantification of Highly-Parameterized Geothermal Reservoir Models Using Ensemble-Based Methods. Available online: <https://pangea.stanford.edu/ERE/db/WGC/Abstract.php?PaperID=5306> (accessed on 23 May 2020).
52. Bezak, N.; Rusjan, S.; Petan, S.; Sodnik, J.; Mikoš, M. Estimation of soil loss by the WATEM/SEDEM model using an automatic parameter estimation procedure. *Environ. Earth Sci.* **2015**, *74*, 5245–5261. [[CrossRef](#)]
53. Doherty, J. Calibration and Uncertainty Analysis for Complex Environmental Models. *Groundwater* **2015**, *53*, 673–674. [[CrossRef](#)]
54. Seibert, J.; Vis, M.J.P.P. Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 3315–3325. [[CrossRef](#)]
55. Seibert, J. HBV-Light Data Exercise Link. Available online: <https://www.geo.uzh.ch/en/units/h2k/Services/HBV-Model/HBV-Download.html> (accessed on 21 January 2020).
56. Doherty, J.; Hunt, R. *Approaches to Highly Parameterized Inversion: A Guide to Using PEST for Groundwater-Model Calibration*; U.S. Geological Survey: Reston, VA, USA, 2010; p. 59. Available online: <https://pubs.usgs.gov/sir/2010/5169/> (accessed on 23 May 2020).

57. Doherty, J. PEST Version 17. Available online: [www.http://pesthhomepage.org/Downloads.php](http://pesthhomepage.org/Downloads.php) (accessed on 21 January 2020).
58. Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [[CrossRef](#)]
59. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **1944**, *2*, 164–168. [[CrossRef](#)]
60. Doherty, J. PEST: Model-independent parameter estimation. *Watermark Comput. Corinda Aust.* **2005**, *2005*, 122.
61. Iskra, I.; Droste, R. Application of Non-Linear Automatic Optimization Techniques for Calibration of HSPF. *Proc. Water Environ. Fed.* **2007**. [[CrossRef](#)]
62. Doherty, J. *PEST\_HP, PEST for Highly Parallelized Computing Environments*; Watermark Numerical Computing: Brisbane, Australia; p. 2020.
63. Gavin, H. The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. *Dep. Civ. Environ. Eng. Duke Univ.* **2011**, 1–15.
64. Razavi, S.; Gupta, H.V. A new framework for comprehensive, robust, and efficient global sensitivity analysis: 2. Application. *Water Resour. Res.* **2016**, *52*, 440–455. [[CrossRef](#)]
65. Kalman, D. A Singularly Valuable Decomposition: The SVD of a Matrix. *Coll. Math. J.* **1996**. [[CrossRef](#)]
66. Tonkin, M.J.; Doherty, J. A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resour. Res.* **2005**, *41*. [[CrossRef](#)]
67. Doherty, J.; Johnston, J.M. Methodologies for calibration and predictive analysis of a watershed model. *J. Am. Water Resour. Assoc.* **2003**, *39*, 251–265. [[CrossRef](#)]
68. Fienen, M.N.; Muffels, C.T.; Hunt, R.J. On constraining pilot point calibration with regularization in PEST. *Ground Water* **2009**, *47*, 835–844. [[CrossRef](#)] [[PubMed](#)]
69. Doherty, J. *PEST Model-Independent Parameter Estimation User Manual Part I: PEST, SENSAN and Global Optimisers*; Watermark Numerical Computing: Brisbane, Australia, 2004.
70. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *NUMERICAL RECIPES The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2007; ISBN 9788578110796. [[CrossRef](#)]
71. Seibert, J. HBV-light Model Program. Available online: <https://www.geo.uzh.ch/dam/jcr:2da100b8-5cc4-4626-aa26-3588c914a6b8/HBV-light.zip> (accessed on 21 January 2010).
72. Konz, M.; Seibert, J. On the value of glacier mass balances for hydrological model calibration. *J. Hydrol.* **2010**, *385*, 238–246. [[CrossRef](#)]
73. Lindström, G.; Johansson, B.; Persson, M.; Gardelin, M.; Bergström, S. Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.* **1997**, *201*, 272–288. [[CrossRef](#)]
74. Gupta, V.K.; Sorooshian, S. Uniqueness and observability of conceptual rainfall-runoff model parameters: The percolation process examined. *Water Resour. Res.* **1983**, *19*, 269–276. [[CrossRef](#)]
75. Renard, B.; Kavetski, D.; Kuczera, G.; Thyer, M.; Franks, S.W. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resour. Res.* **2010**, *46*. [[CrossRef](#)]
76. Doherty, J. The HP Suite. Available online: [www.http://pesthhomepage.org/Downloads.php](http://pesthhomepage.org/Downloads.php) (accessed on 21 January 2020).
77. Pool, S.; Vis, M.; Seibert, J. Evaluating model performance: Towards a non-parametric variant of the Kling-Gupta efficiency. *Hydrol. Sci. J.* **2018**, *63*, 13–14. [[CrossRef](#)]
78. Smakhtin, V.U. Low flow hydrology: A review. *J. Hydrol.* **2001**, *240*, 147–186. [[CrossRef](#)]
79. Velázquez, J.A.; Schmid, J.; Ricard, S.; Muerth, M.J.; Gauvin St-Denis, B.; Minville, M.; Chaumont, D.; Caya, D.; Ludwig, R.; Turcotte, R. An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 565–578. [[CrossRef](#)]
80. Dams, J.; Nossent, J.; Senbeta, T.B.; Willems, P.; Batelaan, O. Multi-model approach to assess the impact of climate change on runoff. *J. Hydrol.* **2015**, *529*, 1601–1616. [[CrossRef](#)]
81. Sapač, K.; Rusjan, S.; Šraj, M. Assessment of consistency of low-flow indices of a hydrogeologically non-homogeneous catchment: A case study of the Ljubljanica river catchment, Slovenia. *J. Hydrol.* **2020**. [[CrossRef](#)]
82. Lindström, G. Lake water levels for calibration of the S-HYPE model. *Hydrol. Res.* **2016**, *47*, 672–682. [[CrossRef](#)]

83. Hottel, C.; Braun, L.N.; Leibundgut, C.; Rieg, A. Simulation of snowpack and discharge in an alpine karst basin. In Proceedings of the International Symposium, Kathmandu, Nepal, 16–21 November 1992.
84. Arsenault, R.; Poulin, A.; Côté, P.; Brisette, F. Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration. *J. Hydrol. Eng.* **2014**, *19*, 1374–1384. [[CrossRef](#)]
85. Brilly, M.; Kryżanowski, A.; Šraj, M.; Bezak, N.; Sapač, K.; Vidmar, A.; Rusjan, S. Historical, Hydrological and Hydraulics Studies for Sustainable Flood Management. In *Achievements and Challenges of Integrated River Basin Management*; IntechOpen: London, UK, 2018. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).